More investment in Research and Development for better Education in the future?

Rim Lahmandi-Ayed and Dhafer Malouche

ESSAI and MASE-ESSAI, University of Carthage

Abstract

The question in this paper is whether R&D efforts affect education performance in small classes. Merging two datasets collected from the PISA studies and the World Development Indicators and using Learning Bayesian Networks, we prove the existence of a statistical causal relationship between investment in R&D of a country and its education performance (PISA scores). We also prove that the effect of R&D on Education is long term as a country has to invest at least 10 years before beginning to improve the level of young pupils.

Keywords: R&D, Education performance, Learning Bayesian Networks, PISA, WDI.

1. Introduction

Education is considered among the basic human rights all over the world. This is not surprising as an abundant literature recognizes the numerous benefits of education. Besides being desirable in itself, education has positive effects at the individual level in terms of wealth, health, happiness within the couple and the family; favors social cohesion reducing crimes and raising voluntary activities; and additionally has positive economic benefits, increasing productivity, economic growth and competitiveness.

Given the importance of education in human life, it seems on the one hand necessary to study the determinants of education performance. On the other hand, numerous studies measure the short-term effects of Research and Development. But "we know surprisingly little about their long-term effects. This is a pity, because the conventional justifications for state intervention in research depend upon phenomena that are inherently long term in nature." (Arnold, 2012) To the best of our knowledge, there is no study linking R&D and education although both are components of the same educative system.

In this paper, we argue that R&D expenditure affects positively education in the long run. More precisely, we prove the existence of a statistical causal relationship between investment in R&D and education performance (PISA scores). We also prove that this effect is long term as a country has to invest at least 10 years before beginning to improve the level of young pupils.

The scarcity of studies of long-term effects of R&D, among which the effect on education performance, is due to the fact that the stakeholders are only concerned by effects expected in a horizon relevant to their elective life or budgeting cycle[1]. But because long term effects of R&D are hardly studied, the decisions taken concerning R&D may be sub-optimal ones, as they do not take those effects into account.

Measuring the effect of R&D expenditure on education performance in small classes and how long we must wait for these effects, is very important in general. It is more important in particular in developing countries where research is not given much importance, believing that they can do without research by benefiting from the research of others through ready-for-use items and technologies, and choosing to focus on education in primary and secondary schools. But investment in R&D affecting education of children, decision-makers must understand that if R&D is neglected, not only researchers and universities will suffer but the performance of primary and secondary schools is also at stake, which may threaten the whole educative system.

Our proof in this paper is based on a statistical modeling performed from merging two datasets collected from two different sources: the PISA studies made by the OECD and the World Development Indicators provided by the World Bank.

"The Programme for International Student Assessment (PISA) is a triennial international survey which aims to evaluate education systems worldwide by testing the skills and knowledge of 15-year-old students" in five disciplines[2].We consider the 2015 study, the most recent one, to benefit from the maximal hindsight, focusing on the national scores obtained in mathematics, reading and science which are our *output* variables.

The second dataset used in this paper is provided by the World Bank

[1] This is the so called "dynamic inconsistency" existing between the long time constants of R&D and shorter cycles in political, administrative and budgeting terms referred to by Arnold (2012).

[2] http://www.oecd.org/pisa/

in the so-called World Development Indicators available on the web and in a package from the R statistical software[3]. This database allows the computation of the annual expenditure per researcher of each country.

We use Learning Bayesian Networks to estimate the relationships in the random vector constituted by the annual R&D expenditure per researcher from 1997 to 2014 and the PISA scores in 2015 (Pearl, 2000; Spirtes and Glymour, 1991; Scutari and Denis, 2015). Before making the estimation, the Bayesian Networks are also used to impute the missing values, thus providing an imputation method better than the existing ones.

Noting that the three considered scores (reading, mathematics and science) are strongly correlated we restrict our study to reading only to avoid redundancy. We then measure the strength of the link between each annual R&D expenditure and the reading score. We prove that the investment the most likely linked to the "Reading" variable is the one made in 2005. This means that the efforts made in R&D begin to show results only 10 years later.

As an interpretation of the obtained results, we think that R&D affects education in small classes through the training of trainers and all types of educators who will be in charge of children including future parents. The money spent in R&D raises the quality of university teaching, thus the quality of the graduates among whom the future adults in charge of children. This explains in the same time why the delay is long. Indeed the effect of R&D on education is not a direct one. It occurs through a chain of effects, each needing some delay. A dollar spent in R&D has first to impact the performance of researchers (publications, patents...), which in turn impacts the quality of university teaching at the individual level and through a richest more dynamic university environment. Better university teachers give better graduates, so better future parents and educators, which finally impacts the education performance.

The delay of 10 years is too long for countries under financial stress and is beyond the horizon relevant to the elective life or budget cycles in democracies. Cutting the research budgets is almost a reflex of decision-makers at each crisis, this must not be surprising. But this reflex is a bad one, an irresponsible decision as it undermines the education of the future generations. What happens in this respect is very similar to environment issues where decision-makers may take decisions for which they do not have to bear the consequences.

[3] r-project.org

As a by-product of the estimation, we measure the contribution of R&D in the performance of Education (Contribution Index) in each country. This index ranges from 45% to 64%, which is very high, meaning that R&D contributes highly in the performance of Education.

We also evaluate the efficiency of R&D in the performance of Education (Efficiency Index) obtained from the comparison of the estimated performance of education using our model and the observed value. This index, varying from -35 % to +15 %, has necessarily something to do with governance issues and R&D organization. Both indices turn out to be highly correlated, which is natural as R&D contributes more in education performance thanks in great part to its efficiency. This is to say that investing more in R&D may be insufficient. Budgets have to be more efficiently spent and R&D more efficiently organized in order to secure the transmission of the benefits to education performance.

The literature review. Two points have to be made clear. First we deal with research in general and not particularly with research in educational sciences as is done by Cooper (1953), Levin (2004) and certainly numerous others. Second, to the best of our knowledge, there is no literature linking R&D and education.

We can relate our paper to an abundant literature, empirical and theoretical, on the socio-economic impact of knowledge in a broad sense: schooling and basic education (Haveman and Wolfe, 1984), Higher education (O'Carroll et al., 2006) and research and development (Beck et al., 2017).

There are numerous studies on economic effects of R&D: productivity (Griliches, 1998 and Lichtenberg and Siegel, 1991), technological change (Schumpeter, 1942) which is the source for sustainable economic growth (Solow, 1956). Because of spillover effects of the knowledge creation between firms (Beck et al., 2017; Hall et al., 2010; Bloom et al., 2013; Cardamone, 2012; Venturini, 2015; Acharya, 2015; Bloch, 2013), the market may fail in providing the optimal level of R&D investment, which justifies public funding of R&D (Martin and Scott, 2000; Romer, 1990; Jones and Williams, 1998).

Numerous papers are interested in the effects of public and university research. Jaffe et al. (1993), Jaffe and Trajtenberg (1996), Mohnen (1996), Blomstrom and Kokko (1998), Cincera and van Pottelsberghe (2001) study the role of university R&D in increasing productivity and improving the competitiveness of countries. Papers identify the importance of academic research in driving economic growth (Huggins and Cooke, 1997) and regional economic development (Smilor et al., 1993). Science parks or innovation centers most commonly located on university campus and facilitating many spillovers and benefits from the proximity with researchers and innovation, help new companies flourish, which is a channel to create jobs and foster economic growth (O'Carroll et al., 2006).

A stream of literature deals with the socio-economic effects of basic education (schooling or basic education completion) and/or higher education (access or completion). The theoretical papers of Barro (1991) and Lucas (1988) and the empirical study of Dension (1985) demonstrate the positive role of education (or more generally human capital formation) in economic growth. Moretti (2004) establishes that an increase in the rate of graduates over the population of workers increases the wages for all workers. Schultz (1961), Hansen (1963), Becker et al. (1964), Mincer (1962) among others have earlier studied the relation between productivity and schooling.

A body of literature deals with non-marketed effects of education. First education may be consumed for its intrinsic value (Lazear, 1977). Several studies exist on the relation between higher education and voluntary activities: Freeman (1997), Vaillancourt (1994), Gibson (2001), Dee (2003); and between education and money and time donations: Dye (1980) and Mueller (1978). Several papers examine the relation between education and health issues: Grossman (1975), Fuchs (1982), Leigh (1981), Farrell and Fuchs (1982), Kenkel (1991) and Lleras-Muney (2005). Ehrlich (1975) has shown that education reduces criminal activity and Lochner and Moretti (2001) investigate the relationship between High School completion and crime. Numerous papers establish a positive correlation between education and attainment of desired family size: Michael (1973), Ryder and Westoff (1965), Michael and Willis (1976); and a positive correlation between education and sorting in the marriage market (Becker et al., 1977 and Jensen, 1969). Importantly for our purpose, numerous papers have proved that mother's and father's education positively influences child quality in several respects (health, cognitive development, education, occupation status, future earnings): Leibowitz (1974, 1975), Edwards and Grossman (1979), Birch and Gussow (1970), Hill and Stafford (1974, 1980), Wolfe and Behrman (1982), among numerous others.

To the best of our knowledge, no paper deals with the effect of R&D expenditure on education performance. In a broader sense, there is no paper dealing with the determinants of education performance. However there are numerous papers dealing with the determinants of research and innovation. Berman (1990) establishes that university-funded research stimulates the industrial R&D. Numerous papers deal with the effects of direct and/or indirect public funding for private R&D on the firms' efforts in R&D: Dimos and Pugh (2016), Hud and Hussinger (2015), Duguet (2004) among numerous others. Science parks and innovation centers foster industry oriented innovations through the transfer of academic research. Finally some old papers backward trace innovations of interest (such as weapons, oral contraceptives...) and identify the scientific events that fed into them: Iserson (1967), Loellbach (1968, 1969).

Finally, our paper may be related to De Pillis and G. (2001) who develop a mathematical model to explore the long-term effects of university funding on the population of active professors, the number of students and the creation of jobs in the industry.

The remainder of the paper is organized as follows. Section 2 describes the data. Section 3 presents the theory used in the imputation and in the estimation of the model. Section 4 provides the results. Section 5 discusses the obtained results. Section 6 concludes.

2. Materials and methods

In this paper we design a dataset from two very well known sources. The first source called PISA data is collected from the Program for International Student Assessment (PISA). This is an international survey which aims to evaluate education systems worldwide by testing the skills and knowledge of 15-year-old students[4]. Every three year period, beginning from 2000, around one half million students from over 72 countries are tested in five skills. In this paper we focus on the performance in three among them: mathematics, reading and science, and we use the most recent PISA study, i.e. the 2015 study, in order to benefit from the maximal hindsight.

The raw data are individual marks obtained by the students in each country. This data is available in either a SAS or SPSS format (`http://www.oecd.org/pisa/data/2015database/`) which are large data files, around 400 MB each. Once downloaded, we use the R package `intsvy` (Caro and Biecek, 2017) to compute the national scores in mathematics, science and reading obtained as the average of the individual marks within each country. Denote by $\mathcal{Y}$ the obtained series of observations of the PISA national scores constituting our *output* variables.

The second data is obtained from a database known as the World Development Indicators (WDI) collected by the World Bank. It consists in a grouping of 800 indicators covering more than 150 economies and compiled from officially recognized national sources. This data presents the most re-

[4] `http://www.oecd.org/pisa/`

cent and accurate global development data available and includes national, regional and global estimates[5]. This data is downloaded and managed with the `R` package `WDI`[6]. The R&D indicators contained in the WDI database are available only from 1997 to 2014 with some reasonable missing values. This is why we limit our study to this period.

To measure the investment of countries in R&D, we choose to consider the annual expenditure per researcher. Indeed the percentage of GDP granted to R&D cannot be the right variable as it varies very little among countries. Neither is the gross amount granted to R&D as it varies too much and must be put in perspective relative to the number of researchers. The annual expenditure per researcher appears to be a reasonable choice with a reasonable variance. We have to calculate it as it is not available directly in the database. To do so, we use the two following indicators among the ones available in the WDI database:

- **Expend**: Research and Development expenditure (% of GDP). It is defined as the amount in dollars of the Expenditures for Research and Development. It includes the current and capital expenditures (both public and private) on creative work undertaken systematically to increase knowledge, including knowledge of humanity, culture, and society, and the use of knowledge for new applications. R&D covers basic research, applied research, and experimental development.
- **NumbRD**: Number of researchers. It is defined as the number of Researchers per million people. The researchers are defined to be the professionals engaged in the conception or creation of new knowledge, products, processes, methods or systems and in the management of the projects concerned, including postgraduate PhD students.

We also extract from the WDI data two other indicators: the GDP in dollars (**GDP**) and the total size of the population (**Pop**). This enables us to compute the two following indicators:

- Total amount of Expenditure in dollars
$$\mathbf{TotExp} = \mathbf{Expend} \times \mathbf{GDP} \times 10^{-2}. \tag{1}$$
- Total Number of Researchers
$$\mathbf{TotRD} = \mathbf{NumbRD} \times \mathbf{Pop} \times 10^{-6}. \tag{2}$$

[5] `https://datacatalog.worldbank.org/dataset/world-development-indicators`
[6] `https://github.com/vincentarelbundock/WDI`

We thus obtain the Expenditure per Researcher

$$\mathbf{ExpOneRD} = \frac{\mathbf{TotExp}}{\mathbf{TotRD}}, \tag{3}$$

and a time series for each country defined by the annual expenditure spent by a given country for the activity of one researcher from year 1997 to year 2014:

$$\mathbf{ExpOneRD}(t), \quad t = 1997, \dots, 2014.$$

It is more appropriate to use a logarithm transformation of the $\mathbf{ExpOneRD}(t)$ variables instead of using them in a raw format. Indeed when we look at the relationship between the **ExpOneRD** variables and the PISA scores we notice that above a certain amount of investment, the PISA scores vary very little. There is indeed a kind of bearing in the scatter plot (see Figure 1) suggesting rather a logarithmic relationship between the **ExpOneRD** variable and the PISA scores.

Hence, our final *input* data, denoted by $\mathcal{X}$ and called the RD investment data, is composed of the variables equal to the logarithm transformation of $\mathbf{ExpOneRD}(t)$: $\mathcal{X} = \{\log(\mathbf{ExpOneRD}(t)),\ t = 1997, \dots, 2014\}$.

The last Data design step consists in merging $\mathcal{X}$ and $\mathcal{Y}$, the two recently constructed datasets, in order to obtain one data containing the PISA national scores of 2015 (either Math or Reading or Science) and the RD investment data $\mathcal{X}$. This restricts our analysis to the 57 countries or economies belonging in the same time to both datasets. Our data can be denoted by

$$\mathcal{D} = [\mathcal{Y}, \mathcal{X}]\,. \tag{4}$$

But because of the missing values in the raw data $\mathcal{X}$, we need to go through a process of data cleaning and imputation. This procedure is somewhat innovative and also based on the estimation of Bayesian Networks as we will explain in the next section.

3. Theory

We have constructed a data $\mathcal{D}$ that can be considered as an $n-$sample of observations of the random vector

$$[Y, \mathbf{X}] = [Y, (\,X_{1997}, \dots, X_{2014})]\,, \tag{5}$$

where Y is a random variable representing one of the PISA scores (Math, Reading or Science) and $X_{1997}, \dots, X_{2014}$ are the random variables representing the $\log(\mathbf{ExpOneRD})$ variables. Our objective in this paper is to estimate the probability distribution of the random vector $[Y, \mathbf{X}]$.

To do so, we estimate the model as a Bayesian Network using Gaussian Bayesian Networks called Learning Bayesian Networks. We first describe

Bayesian Networks, then the estimation process and finally the imputation method also using Bayesian Networks.

Bayesian Networks (BN) are Directed Acyclic Graphs (DAG) used to read the relationships between the variables in the random vector $[Y, \mathbf{X}]$. Mathematically speaking a BN is a couple $G = (V, E)$ where V is the set of nodes. Each node $v \in V$ represents one variable from the random variables associated to the input or output data $\{Y\} \cup \{X_t, t = 1997, \ldots, 2014\}$, and E is the set of directed edges.

E is a subset of $V \times V$ such that, if $(v, v') \in E$ then $(v', v) \notin E$. An element (v, v') in E is then denoted by $(v \to v')$. Thus the elements of E are arrows with directed edges. We denote for every $v \in V$ by $\theta(v)$ the variable in $[Y, \mathbf{X}]$ represented by the node v in the DAG G. $\theta(v)$ can be either Y (the PISA score) or one of the variables X_t (the investment variables in R&D). If $S \subseteq V$, the sub-random vector indexed by S in $[Y, \mathbf{X}]$ will be denoted by $\Theta(S) = (\theta(s), s \in S)$.

If $(v \to v') \in E$ then v is called the *parent* of v' and v' the *child* of v. A *path* from an edge v to another edge v' is a sequence of edges $v_0, \ldots, v_n$ such that $v_0 = v$, $v_n = v'$ and for every $i = 0, \ldots, n-1$, $(v_i \to v_{i+1}) \in E$.

The DAG G, when associated to a random vector $[Y, \mathbf{X}]$, is used to read the conditional independence between the variables in the random vector $[Y, \mathbf{X}]$. In fact G helps to give a parsimonious factorization of the probability density of the random vector $[Y, \mathbf{X}]$. Hence if f is the density of $[Y, \mathbf{X}]$, this density may factorize when G is known, i.e. may take the following form:

$$f(\Theta) = \prod_{v \in V} g(\theta(v) \mid \Theta(pa(v))), \tag{6}$$

where $\Theta = (\theta(v), v \in V) \in \mathbb{R}^{|V|}$ is the random vector $[Y, \mathbf{X}]$ and $pa(v)$ denotes the set of parents of the node v in G.

Equality (6) is called the factorization Markov property of the probability of $[Y, \mathbf{X}]$ w.r.t. G. When (6) is satisfied we can also conclude that for any couple of nodes v and v' the following pairwise Markov property is satisfied:

$$\text{If neither } (v \to v') \text{ nor } (v \to v') \in E \;\text{ then }\; \theta(v) \perp\!\!\!\perp \theta(v') \mid \Theta(pa(v)), \tag{7}$$

where $\theta(v) \perp\!\!\!\perp \theta(v') \mid \Theta(pa(v))$ means that the variables $\theta(v)$ and $\theta(v')$ are independent given the variables of the random vector $\Theta(pa(v))$.

Hence from a statistical point of view when the data is generated from variables satisfying the Factorization Markov Property as in (6) the variables represented by the nodes not belonging to the set of parents of v have no impact on the variable $\theta(v)$. This motivates our choice of BN to study the effect of the investment in R&D on Education. By estimating the BN from the data $\mathcal{D}$ we can not only know which year's investment has a direct impact on the PISA scores but we also know the whole mechanism through which investments in R&D affect the PISA scores. These relationships can also be easily visualized using the representation of the DAG.

3.1. Learning Bayesian Network

To achieve our objective we need to estimate from our collected data $\mathcal{D}$ the Bayesian Network G as defined above. Since this data is exclusively composed of the observation of quantitative variables, we use Gaussian Bayesian Network to model the relationship between the PISA score Y and the investment variables X_t, $t = 1997$ to $t = 2014$. This process of estimation of G is usually called the Learning structure of a Bayesian Network. Several algorithms exist in the literature for this purpose. We usually estimate the BN that corresponds to the minimum of a score measuring how good the BN fits the data. The learning procedure is then a search of a minimum in a discrete space which is the set of possible DAGs. This problem is known to be an NP-hard problem and we have to use some searching algorithm adapted to the BN.

First notice that we are not searching in the whole space of possible DAGs with nodes in V. In fact our DAGs should not contain edges from $X_{t'}$ to X_t when $t' > t$, or edges from Y to any of the X_t when t and t' belong to $\{1997, \dots, 2014\}$, as a variable at some year (investment in R&D or the PISA scores in 2015) is not expected to influence the past. Hence from the investment variable in year 1997 we can have 18 possible arrows or edges, i.e. 17 years from 1998 to 2014 and the Education variable, the variable 1998 will have 17 edges, and so on. So the number of the possible BN fitting our problem is:

$$18! = 18 \times 17 \times \ldots \times 1 = 6.402374 \times 10^{15},$$

which remains a huge number.

This is why sampling algorithms are necessary in this kind of problems. In the literature, there are mainly two families of *learning* Bayesian networks: Constraint-based Algorithms[7] and Score-based Algorithms[8].

But there exist also hybrid algorithms which are a composition between constraint-based and score-based algorithms. We obtain algorithms like the Sparse Candidate algorithm (SC) (Friedman et al., 1999) and the Max-Min Hill-Climbing algorithm (MMHC) (Tsamardinos et al., 2006).

[7]for example the PC-algorithm Spirtes and Glymour, 1991 and Kalisch and Bühlmann,2007)

[8]One of them is the *greedy search algorithms* such as Hill-Climbing with random restarts (Bouckaert, 1995)

In this paper we choose to use the Hill-Climbing (HC) and MMHC algorithms since they are already implemented in the `R` package `bnlearn`[9] (Scutari and Denis, 2015).

The Hill-Climbing (HC) algorithm will provide at the end an estimation of the BN without any measure of the accuracy of this estimation. To obtain a measure of this accuracy we use a Bootstrap procedure (Efron and Tibshirani, 1993). This is well known technique in statistics allows assigning measures of accuracy and provides an estimation of the sampling distribution of any statistics. Since we are here interested by the estimation of the presence or not of directed arrows between pairs of nodes of V, by applying Bootstrap procedure we will be able to give a probability of the presence of an edge in the searched DAGs. The use of Bootstrap procedure in the estimation of BN is already implemented in `R` using the command `boot.strength` of the package `bnlearn`. We use 500 of bootstrap replicates and use for every sample the Hill-Climbing algorithm. We obtain the strength of the presence of each edge among which we are especially interested in the ones linked to the Education variable Y. First we have to see if this probability is high and then we consider the highest probability as a threshold in order to compute an average model containing only the significant edges.

These procedures can be implemented only on complete data. However the investment data $\mathcal{X}$ has missing values. Hence we have first to go through a missing value imputation procedure to complete the missing observations before estimating our model. This is the object of the following subsection.

3.2. Imputing missing values

The investment data $\mathcal{X}$ described above and collected from the WDI website contains some missing values due to non-collected data. In the data $\mathcal{X}$ we have about 205 missing observations among 1057 ones. The years 1997 and 1998 have the highest numbers of missing values. It is about 19 and 20 missing values among the 57 countries. Algeria, Albania and Vietnam are the countries who have provided the highest numbers of missing values. Figure 2 provides all the information on missing values.

Many techniques in statistics exist to impute missing values in multivariate data but there exists one method based on the estimation of Bayesian Networks and it is implemented in the command `impute` in the package `bnlearn`. This technique works like a prediction of new observations by using the observations of the parents of the missing observations, the esti-

[9] `http://www.bnlearn.com`

mated BN and its fitted parameters. This method seems suitable in our case study but at this stage we have not yet estimated the BN and we can not do so since the estimation of the BN precisely requires the dataset without any missing values. This problem looks like the egg-and-chicken problem.

To solve the problem we set up a new procedure based on an iterative method giving an imputation in our opinion better than the existing imputation methods and we call the obtained algorithm Bayesian Network Iterative Imputation algorithm (BNII). We mimic the famous cross-validation method used in classification and regression methods to provide an estimation of the predicted error. At each step of the iteration, we drop randomly real observations, make them missing, impute them using BN and measure the gap between the new imputed values and their real observed values as defined in (8). We stop the procedure when we reach a reasonable low value of this gap. In practice we have run this procedure from 100 to 2000 times by considering steps of 50 iterations. In each iteration we record the minimum of the gap D defined in (8) and its rank in the iteration. We summarize this result in Figure 3, where we put on the $x-$axis the number of iterations and D on the $y-$axis and where the annotated integer is the rank of the observed minimum at the corresponding iteration.

We can notice that D increases slowly when the number of iterations increases and its minimum can be reached before the end of the fixed number of iterations. We observe that this minimum is reached when the procedure is run 1900 times and the minimum is observed at the 638-th iteration.

A necessary initial step consists in an imputation using the KNearest Neighbors algorithm (KNN). It consists in replacing the missing value by the median of the 10 closest neighbors. We thus obtain an initial imputed data $\mathcal{X}_0^c$.

Then we start our iterative procedure. First the BN is estimated from the completed data from the previous step. Then we remove randomly $n = 50$ observations. Using the fitted model we estimate the missing values to complete the data and measure D. We continue this iteration for a fixed number of times. The whole procedure is described in Algorithm 1 where

- $\mathcal{S}_T$ is the whole set of indices of the observations in the data $\mathcal{X}$, i.e. $\mathcal{S}_T = \{1, \ldots, n\} \times \{1, \ldots, p\}$,
- n and p are respectively the number of rows and columns in $\mathcal{X}$,
- $\mathcal{S}_{NA}$ is the set of indices of the missing observations in the data $\mathcal{X}$,
$$\mathcal{S}_{NA} \subset \mathcal{S}_T.$$

Algorithm 1 Bayesian Network Iterative Imputation Algorithm

let $i = 0$.
Start with $\mathcal{X}_i$ a complete version of $\mathcal{X}$ computed using KNN algorithm.
let N be the maximal number of iterations for the coming loop.
Let fix $d_i > 0$ an initial large positive number.
while $i \leqslant N$ **do**
    Sample from $\mathcal{S}_T \backslash \mathcal{S}_{NA}$ 50 indices complete observations from $\mathcal{X}$ and transform them to missing values. Denote by $\mathcal{S}_r$ this set of indexes, i.e; $\mathcal{S}_r \subset \mathcal{S}_T \backslash \mathcal{S}_{NA}$
    Denote by $\mathcal{X}'$ the new version of $\mathcal{X}$ with missing values at $\mathcal{S}_r \cup \mathcal{S}_{NA}$.
    Estimate and fit the BN using HC algorithm from $\mathcal{X}_i$. Let's denote by $\widehat{G}$ the estimated BN model.
    Impute $\mathcal{X}'$ using $\widehat{G}$ and obtain a new complete version of $\mathcal{X}$. Let's denote it by $\mathcal{X}_c$.
    Compute

$$D = \sum_{s \in \mathcal{S}_r} (x(s) - x_c(s))^2 \tag{8}$$

    where $x(s)$ and $x_c(s)$ are respectively is the generic coefficient of $\mathcal{X}$ and $\mathcal{X}_c$.
    if $d_i \geqslant D$ **then**
        $i \leftarrow i + 1$, $d_i \leftarrow D$ and $\mathcal{X}_i = \mathcal{X}_c$.
    else
        $i \leftarrow N + 1$
return $\mathcal{X}_i$ a complete version of the data

4. Results

Figures 4 and 5 are our first outputs. In Figure 4 we represent the countries with respect to their three PISA Scores: Math, Reading and Science, sorting them from the less performing, Algeria and Tunisia to the most performing countries, Singapore, Hong Kong, Macao and Japan. We notice from this graph the strong correlation between these three scores and we can even conclude that they are redundant. This is why we have decided to consider only the "Reading" score to understand the impact of the investment in R&D on the performance of Education.

Figure 5 is a representation of the correlation of the PISA scores with each of the investment variables in R&D from 1997 to 2014. We also represent with segments the 95% confidence interval (CI) of each correlation coefficient. First we notice that none of these CIs contains zero. This means that the relationship between investment in R&D and performance in Education is always significant. Hence whenever you invest in R&D there is always a significant impact on the performance of Education. One important question is when this investment has the highest impact on Education, i.e. is the effect stronger in the short or the long run? We indeed use a sophisticated model such as the Bayesian Networks to make sure that we are computing properly the effects of the investments in R&D of all the preceding years on Education and be able to compare them. The years from 2002 to 2005 seem to have slightly higher correlation than all the other years. We will next check that the estimation of Bayesian Network confirms this preliminary observation.

We run 500 times the bootstrap procedure to estimate the Bayesian Network. We represent in Figure 6 the strength of each arrow pointing from each investment year to the Education variable "Reading". We can easily notice that the investment the most likely linked to the "Reading" variable is the one made in 2005. For example the arrow linking the variables representing the investment in R&D in 2005 and "Reading" is present in 70% of the estimated BN. The arrow pointing from R&D in 2014 to "Reading" is only present in 14% of the estimated BN amid the 500 bootstrap replicates. This fact shows obviously that the highest impact on the performance of Education is recorded from the investment in R&D 10 years earlier.

We have chosen a threshold equal to 60% when computing the average of the models among the 500 estimated during the bootstrap procedure. The estimated BN is then represented in Figure 7. From this BN we can notice that the investment in 1997 affects the one in 1998 and 1999 and this impact will be successively continued until 2005 which is the last year that impacts highly the performance "Reading".

We have also estimated the parameters in the models by successively performing a series of linear regression models according to the path from the investment variable in 1997 to the PISA scores. This path is colored in gray in the representation of the estimated final BN in Figure 7. These estimations are displayed in Table 1. In this table we show the estimation of the parameters of each model determined by each node in the path between the variable representing the investment in 1997 to the "Reading" PISA score. If v is a node in that path, the estimated regression shown in Table 1 is the one defined by $\theta(v)$ as a dependent variable and $\Theta(u)$ as the vector of independent variables, u is a node parent of v, i.e; $(\Theta(u), u \in pa(v))$. Note that in Table 1, the dependent variables are on the columns and the independent are in the rows. Empty cells in Table 1 show that the corresponding variable is not a parent of the dependent variable in the estimated BN. Hence by performing this series of regression models we can deduce the conditional probability distribution of $\theta(v) \mid \Theta(u), u \in pa(v)$ for every v in the path between the variable investment in R&D in 1997 to the variable "Reading". Thus by using the Markov property introduced in (6) we can deduce the estimation of the parameters of the probability distribution of the random vector composed by the variables "Reading" and those on the path between it and the variable R&D in 1997.

5. Discussion

The question is why there is a causal relationship between expenditure in R&D and education in small classes and why such a long delay is necessary.

R&D is intended to better understand the world and make discoveries meant to improve human life some way or the other. This is how most people perceive R&D. But R&D is also intended to train the trainers, those who will be university teachers (through the supervision of Phd theses) but also the teachers who will educate the children in the primary and secondary schools, the future parents and all types of educators who are likely to take care of the children for diverse activities (sports, excursions, cultural and recreational activities...). One dollar spent in R&D has repercussions in the short term on the output of research (publications, patents...) and thus on the level of the university teachers. But the latter has an effect on the quality of training of the students among whom the future parents and the educators of children. This is in line with the numerous papers who have proved that mother's and father's education positively influences child quality in several respects (health, cognitive development, **education**, occupation status, future earnings): Leibowitz (1974, 1975), Edwards and Grossman (1979), Birch and Gussow (1970), Hill and Stafford (1974, 1980), Wolfe and Behrman (1982), among numerous others.

This is not to say that good researchers are necessarily good university teachers or that university teachers cannot be good if they do not do enough research. First it is a *statistical* causal relationship which may fail to hold in some cases. Second, it is the environment of teaching as a whole which is important in universities. When there is a dynamic research in a university, even the teachers who, strictly speaking, do not do research (not writing articles or filing patents for instance) benefit from the "research environment" where they attend seminars, discuss with colleagues who are researchers, meet invited professors thanks to the activity of researchers, are supervised and/or surrounded by active researchers... The training plans evolve better under the influence of university teachers "up-to-date" with the recent discoveries and theories, including pedagogical innovations. The research creates an environment favorable to questioning, competition and progress, which is beneficial to teaching practices and all university teachers even those who are not active researchers. This explanation is confirmed by our results. Indeed the estimated lag between the investment in R&D and the improvement of education performance (10 years) corresponds to the time necessary for the research performance itself to be positively affected by the investment in R&D, then for this performance to affect positively the quality of teaching and finally to improve the education performance through a better quality of graduates. Research boosts this way the universities and the whole educative system.

Two other indices are estimated using the BN model to shed another light on the relation between R&D and Education.

First the contribution of R&D in the performance of Education (Contribution Index) is computed using the estimation of the intercept of regression model where Y is the variable "Reading" as a dependent variable and where $\theta(2005)$ is the variable investment in 2005 as an independent variable. The regression model can be written as

$$Y = \alpha + \beta \times \theta(2005) + \epsilon, \tag{9}$$

where ϵ is the random error supposed to follow a central Gaussian distribution with unknown variance σ^2 and where α and β are respectively the intercept and the slope of the regression. According to the result displayed in Table 1, the estimation of α is equal to $\widehat{\alpha} = 192.574$ and the estimation

Table 1: Regression models from the best fitted BN

	Dependent variable:							
	‘1998‘	‘1999‘	‘2000‘	‘2001‘	‘2003‘	‘2004‘	‘2005‘	Read
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
‘1997‘	0.956***	−0.406***						
	(0.015)	(0.073)						
‘1998‘		1.445***						
		(0.075)						
‘1999‘			0.971***					
			(0.014)					
‘2000‘				0.987***				
				(0.014)				
‘2002‘					0.980***			
					(0.017)			
‘2003‘						0.952***		
						(0.013)		
‘2004‘							0.915***	
							(0.033)	
‘2005‘								25.139***
								(4.620)
Const.	0.481***	−0.448***	0.275*	0.125	0.372**	0.672***	0.984***	192.574***
	(0.159)	(0.096)	(0.148)	(0.149)	(0.184)	(0.147)	(0.366)	(51.807)
Obs.	57	57	57	57	57	57	57	57
R^2	0.987	0.996	0.989	0.990	0.984	0.989	0.934	0.350
Adj. R^2	0.987	0.996	0.989	0.989	0.984	0.989	0.933	0.338
RSE	0.138	0.077	0.128	0.126	0.153	0.119	0.281	37.542
df	55	54	55	55	55	55	55	55
F Stat.	4,343.525***	7,412.672***	5,135.280***	5,225.948***	3,377.611***	5,126.453***	782.197***	29.604***
df	1; 55	2; 54	1; 55	1; 55	1; 55	1; 55	1; 55	1; 55

Note: *p<0.1; **p<0.05; ***p<0.01

of β is equal to $\widehat{\beta} = 25.239$.

The first coefficient $\widehat{\alpha}$ can be interpreted as the expected Education score when no investment is done in R&D. The second coefficient $\widehat{\beta} = 25.239$ can be interpreted as the gain in Education score when it is decided to invest 1% more in R&D. We can conclude that an Education score can gain an average of 25 points when 1% more is invested in 2005. The coefficient $\widehat{\alpha}$ can also be interpreted as the average score in "Reading" that can not be explained by the investment in R&D. We then compute the percentage of performance in Education that can be explained by the investment in R&D (and call it simply Contribution of R&D) as follows:

$$\text{Contribution Index}(w) = \frac{Y(w) - \widehat{\alpha}}{Y(w)}, \tag{10}$$

where, for every country w, $Y(w)$ is its "Reading" score.

The Contribution Index is computed for every country and depicted in Figure 8. Countries are ranked increasingly with the obtained Contribution of R&D. The contribution of R&D in the performance of Education is the weakest in Algeria, Macedonia and Tunisia which also are the last countries in terms of PISA scores.

The contribution index ranges from 45% to 68 %, which is high. This means that R&D contributes generally much in the education performance, even for countries spending little on R&D and performing badly in education. This confirms our main message, i.e. R&D is important to foster education.

A second index is estimated: the efficiency of R&D in the performance of Education (Efficiency Index). This index is a consequence of the comparison for each country between the estimation of the expected PISA score and the observed one. We define for each country w, the Efficiency of R&D in Education performance as the index EffiRD(w) given by the following:

$$\text{Efficiency Index}(w) = \frac{Y(w) - \widehat{Y}(w)}{Y(w)}, \tag{11}$$

where $\widehat{Y}(w)$ is the expected score computed using the adjusted BN model and $Y(w)$ is the observed score.

This index, given by the difference between the observed score $Y(w)$ and the expected Education $\widehat{Y}(w)$ divided by the observed score $Y(w)$, can be interpreted as the Efficiency of the investment of R&D in the Education performance, as it shows whether the considered country is more or less efficient than expected and measures this efficiency or lack of efficiency in percentage of the observed performance. When this index is positive it means that we do better than what is expected and when it is negative we do worse. Figure 9 representing this index shows that the countries having the worst Education scores have also the least Efficient R&D in Education.

We note substantial differences across countries, efficiency of R&D ranging from -35% to +15%. For example, the Education performance of Tunisia is less by 21% than the expected one. But Estonia does 15% better than what it is expected. Among the 57 considered countries, 23 have negative values of Efficiency of R&D in Education performance, suggesting that they can do more with the same invested amounts, or equivalently, that their educative system is somehow deprived from “free” additional performance. Efficiency has necessarily something to do with governance of R&D and of education, a better governance leading to a better performance of education for the same amount spent on Research.

Now considering together the two new indices (Contribution and Efficiency Indices), in Figure 10 countries are depicted with their national PISA scores on the (Contribution, Efficiency)-axes. We note first that the two new indices are highly correlated. This is natural, as a more efficient research is more likely to contribute more in education. But this correlation is not perfect, i.e. the two indices do not measure the same thing. Indeed countries are not ranked in the same order in terms of each index (Figure 11). The differences between countries in terms of Contribution Index are high, partly because countries differ in terms of efficiency and partly because of other variables not accounted for in our model, such as direct investments in education, transportation infrastructure, economic performance, social policy... To conclude, investing more in R&D is important but may be insufficient. One has also to be efficient to perform better with the same amounts. Even if it is not explicit in our modeling, the organization of R&D (thus governance) and the mechanism of transmission of its effects to university teaching and the quality of graduation appear to be important to obtain the desired improvement in education performance.

6. Conclusion

Considering the World Development Indicators and the national PISA scores obtained in 2015 and using Learning Bayesian Networks, we prove that R&D expenditures affect positively education performance, explaining more than 45% of the education performance.

Research contributes highly to boost the universities, we already knew that. Our paper provides evidence that it also boosts the whole educative system. Research is not a luxury devoted to developed countries, and may even be more vital for developing countries where the educative systems may be more fragile. Nor is it a luxury limited to fat years in developed countries. And claiming every now and then, especially in lean years, that researchers do not make enough discoveries, is not an acceptable reason to justify to cut research budgets.

But this conclusion does not absolve countries and researchers from the obligation of trying to make R&D more efficient. Indeed, when measuring the efficiency of R&D in the performance of R&D, we also noted substantial differences between countries, suggesting that 23 among the 57 countries can do better with the same invested amounts. A better governance and organization of research and more efforts for the transmission of the benefits of R&D to university teaching and graduation, are as necessary as investments in R&D, in order to secure the desired improvement of education performance.

Bibliography

[1] E. Arnold, Understanding long-term impacts of rd funding: The eu framework programme, Research Evaluation 21 (2012) 332–343.

[2] J. Pearl, Causality, Cambridge University Press, 2000.

[3] P. Spirtes, C. Glymour, An algorithm for fast recovery of sparse causal graphs, Social Science Computer Review 9 (1991) 62–72.

[4] M. Scutari, J. B. Denis, Bayesian Networks: With Examples in R, Chapman and Hall, CRC, 2nd edition, 2015.

[5] D. H. Cooper, The impact of research on education, The Phi Delta Kappan 35 (1953) 16–20.

[6] B. Levin, Making research matter more., Education Policy Analysis Archives 12(56) (2004).

[7] R. H. Haveman, B. L. Wolfe, Schooling and economic well-being: The role of nonmarket effects, The Journal of Human Resources 19 (1984) 377–407.

[8] C. O'Carroll, C. Harmon, L. Farrell, The economic and social impact of higher education, Irish Universities Association, 2006.

[9] M. Beck, J. M., U. Kaiser, On the effects of research and development: a literature review, DEA (2017).

[10] Z. Griliches, Issues in assessing the contribution of research and development to productivity growth, Bell J. Econ (1979) 92–116.

[11] F. Lichtenberg, D. Siegel, The impact of r&d investment on productivity new evidence using linked r&d-lrd data, Economic Inquiry 29(2) (1991) 203–229.

[12] J. Schumpeter, Creative destruction, Capital. Social. Democr (1942).

[13] R. M. Solow, A contribution to the theory of economic growth, The Quarterly Journal of Economics No. 1 , Vol. 70 (Feb., 1956) 65–94.

[14] B. H. Hall, J. Mairesse, P. Mohnen, Measuring the returns to r&d, Handb. Econ 2 (2010) 1033–1082.

[15] N. Bloom, M. Schankerman, J. Van Reenen, Identifying technology spillovers and product market rivalry, Econometrica 81 (2013) 1347–1393.

[16] P. Cardamone, A micro-econometric analysis of the role of r&d spillovers using a nonlinear translog specification, J. Product 37 (2012) 41–58.

[17] F. Venturini, The modern drivers of productivity, Res. Policy 44 (2015) 357369.

[18] R. C. Acharya, Revisiting measure of r&d spillovers: empirical evidence on oecd countries and industries, Econ 24 (2015) 360–400.

[19] C. Bloch, R&d spillovers and productivity: an analysis of geographical and technological dimensions. econ. innov, New Technol. 8599 (2013) 447–460.

[20] S. Martin, J. Scott, The nature of innovation market failure and the design of public support for private innovation, Res. Policy 437447 (2000) 29.

[21] P. Romer, Endogenous technological change, J. Polit. Econ. 71-102 (1990) 98.

[22] C. I. Jones, J. C. Williams, Measuring the social return to r&d, Q. J. Econ (1998) 1119–1135.

[23] A. B. Jaffe, M. Trajtenberg, R. Henderson, Geographic localization of knowledge spillovers as evidenced by patent citations, in: The Quarterly Journal of Economics, , vol. 108(3), 1993, pp. 577–98.

[24] A. Jaffe, M. Trajtenberg, Flows of knowledge from universities and federal labs: Modeling the flowof patent citations over time and across institutional and geographic boundari, NBER Working Paper No. 5712, National Bureau of Economic Research, Inc. (1996).

[25] P. Mohnen, R&d externalities and productivity growth, STI Review, OECD 39-66 (1996) 18.

[26] M. Blomstrom, A. Kokko, Multinational corporations and spillovers, Journal of Economic Surveys, Blackwell Publishing 12 (1998) 247–77.

[27] M. Cincera, B. van Pottelsberghe, International r&d spillovers: a survey, Cahiers Economiques De Bruxelles 169 (2001) 3–31.

[28] R. Huggins, P. Cooke, The economic impact of cardiff university: Innovation, learning and job creation, GeoJournal 41 (1997) 325–337.

[29] R. Smilor, G. Dietrich, D. Gibson, The entrepreneurial university: The role of higher education in the united states in technology commercialization and economic development, International Social Science Journal 45(1) (1993) 1–11.

[30] R. Barro, Economic growth in a cross section of countries, Quarterly Journal of Economics 106 (1991) 407–502.

[31] R. Lucas, On the mechanics of economic development, Journal of Monetary Economics 22 (1988) 3–29.

[32] E. F. Dension, Trends in American Economic Growth, 1929-1982, Brookings, 1985.

[33] E. Moretti, Estimating the social return to higher education: evidence from longitudinal and repeated cross-sectional data, Journal of Econometrics 121 (2004) 175–212.

[34] T. Schultz, Investment in human capital, The American Economic Review 1(2) (1961) 1–17.

[35] L. Hansen, Total and private rates of return to investment in schooling, Journal of Political Economy 71 (1963) 128–40.

[36] G. S. Becker, H. C. A. Theoretical, E. N. York:, Columbia university press (for nber) (1964).

[37] J. Mincer, Job training: Costs, returns and some implications., Journal of Political Economy 70 (1962) 50–70.

[38] E. Lazear, Education: Consumption or production?, Journal of Political Economy 85 (1977) 569–97.

[39] R. B. Freeman, Working for nothing: the supply of volunteer labour, Journal of Labour Economics S (1997) 140–S166.

[40] F. Vaillancourt, To volunteer or not: Canada, 1987, Canadian Journal of Economics 813-826 (1994) 27.

[41] J. Gibson, Unobservable family effects and the apparent external benefits of education, Economics of Education Review 20 (2001) 225–233.

[42] T. S. Dee, Are there civic returns to education?, Working Paper, NBER, 2003.

[43] R. F. Dye, Contributions to volunteer time: Some evidence on income tax effect, National Tax Journal 33 (1980).

[44] W. Mueller, An economic theory of volunteer work, Department of Economics, Wesleyan University, Middletown, CT, Mimeo (1978).

[45] M. Grossman, The correlation between health and schooling, in: H. P. and (Ed.), and Consumption, ed, National Bureau of Economic Research, N. E. Terleckyj. New York, 1975.

[46] V. R. Fuchs, Time preference and health: an exploratory study, in: R. Fuchs (Ed.), V, Economics Aspects of Health, University of Chicago Press, 1982, pp. 93–120.

[47] P. Leigh, Hazardous occupations, illness and schooling, Economics of Education Review (1981) 381–388.

[48] P. Farrell, V. R. Fuchs, Schooling and health: the cigarette connection, Journal of Health Economics, pp (1982) 217–230.

[49] D. S. Kenkel, Health behavior, health knowledge, and schooling, The Journal of Political Economy 99 (1991) 287–305.

[50] A. Lleras-Muney, The relationship between education and adult mortality in the united states, Review of Economic Studies 72 (2005) 189–221.

[51] I. Ehrlich, On the relation between education and crime, in: Income, and Human Behavior, ed, McGraw-Hill, F. Thomas Juster. New York, 1975.

[52] L. Lochner, E. Moretti, The effects of education on crime: evidence from prison inmates, arrests and self-reports, NBER Working Paper No. 8605, National Bureau of Economic Research, Inc. (2001).

[53] R. Michael, Education and the derived demand for children, Journal of Political Economy 81 (1973) 128–64.

[54] N. Ryder, C. F. Westoff, Reproduction in the United States,, Princeton, NJ: Princeton University Press,, 1965.

[55] R. T. Michael, R. J. Willis, Contraception and fertility: Household production under uncertainty, in: H. P. and (Ed.), and Consumption, ed, National Bureau of Economic Research, Studies in Income and Wealth No. 40, N. E. Terleckyj. New York, 1976.

[56] G. S. Becker, E. M. Landes, R. T. Michael, An economic analysis of marital instability, Journal of Political Economy 85 (1977) 1141–88.

[57] A. Jensen, How much can we boost i.q. and achievement?, Harvard Education Review 39 (1969) 1–123.

[58] A. Leibowitz, Education and the allocation of women's time, in: Income, and Human Behavior, ed, McGraw- Hill, F. Thomas Juster. New York, 1975.

[59] A. Leibowitz, Human investments in children, Journal of Political Economy 82 (1974).

[60] L. N. Edwards, M. Grossman, The relationship between children's health and intellectual development, in: S. Mushkin (Ed.), Health: What Is It Worth?, Pergamon Press, Elmsford, NY, 1979.

[61] H. Birch, J. D. Gussow, Disadvantaged Children: Health, Nutrition, and School Failure, Harcourt, Brace and World, New York, 1970.

[62] R. Hill, F. P. Stafford, Allocation of time to preschool children and economic opportunity, Journal of Human Resources (Summer: (1974) 323–46.

[63] R. Hill, F. P. Stafford, Parental care of children: Time diary estimates of quantity, predictability, and variety, Journal of Human Resources 15 (1980) 219–39.

[64] B. L. Wolfe, J. R. Behrman, Determinants of child mortality, health and nutrition in a developing country, Journal of Development Economics 11 (1982) 163–94.

[65] C. Dimos, G. Pugh, The effectiveness of r&d subsidies: A meta-regression analysis of the evaluation literature, Res 45 (2016) 797–815.

[66] M. Hud, K. Hussinger, The impact of r&d subsidies during the crisis, Res 44 (2015) 1844–1855.

[67] E. Duguet, Are r&d subsidies a substitute or a complement to privately funded r&d ?, Rev. Déconomie Polit. 114 (2004) 245–274.

[68] R. S. Iserson, Project Hindsight (Final Report), Department of Defense, Office of the Director of Defense Research and Engineering, Washington, D.C., 1967.

[69] H. Loellbach (Ed.), Technology in Retrospect and Critical Events in Sciences (TRACES), volume I, Illinois Institute of Technology Research Institute, Contract NSF-C535 with the National Science Foundation, Chicago, 1968.

[70] H. Loellbach (Ed.), Technology in Retrospect and Critical Events in Science (TRACES), volume 2, Illinois Institute of Technology Research Institute, Contract NSF-C535 with the National Science Foundation, Chicago, 1969.

[71] E. G. De Pillis, D. P. L. G., The long-term impact of university budgets cuts: A mathematical model, Mathematical and Computer Modeling 33 (2001) 851–876.

[72] D. Caro, P. Biecek, intsvy: An r package for analyzing international large-scale assessment data, Journal of Statistical Software, Articles 81 (2017) 1–44.

[73] M. Kalisch, P. Bühlmann, Estimating high-dimensional directed acyclic graphs with the pc-algorithm, JMLR 8 (2007) 613–636.

[74] R. R. Bouckaert, Bayesian belief networks: From construction to inference, PhD thesis, Utrecht University, The Netherlands (1995).

[75] N. Friedman, D. Peer, , I. Nachman, Learning bayesian network structure from massive datasets: The "parse candidate" algorithm., Proceedings of 15th Conference on Uncertainty in Artificial Intelligence, (1999) 206221.

[76] I. Tsamardinos, L. E. Brown, C. F. Aliferis, The max-min hill-climbing bayesian network structure learning algorithm, Machine Learning 65(1) (2006) 31–78.

[77] B. Efron, R. Tibshirani, An Introduction to the Bootstrap., Chapman & Hall/CRC., 1993.

[78] L. N. Edwards, M. Grossman, The relationship between children's health and intellectual development, in: S. Mushkin (Ed.), Health: What Is It Worth?, Pergamon Press, Elmsford, NY, 1979.

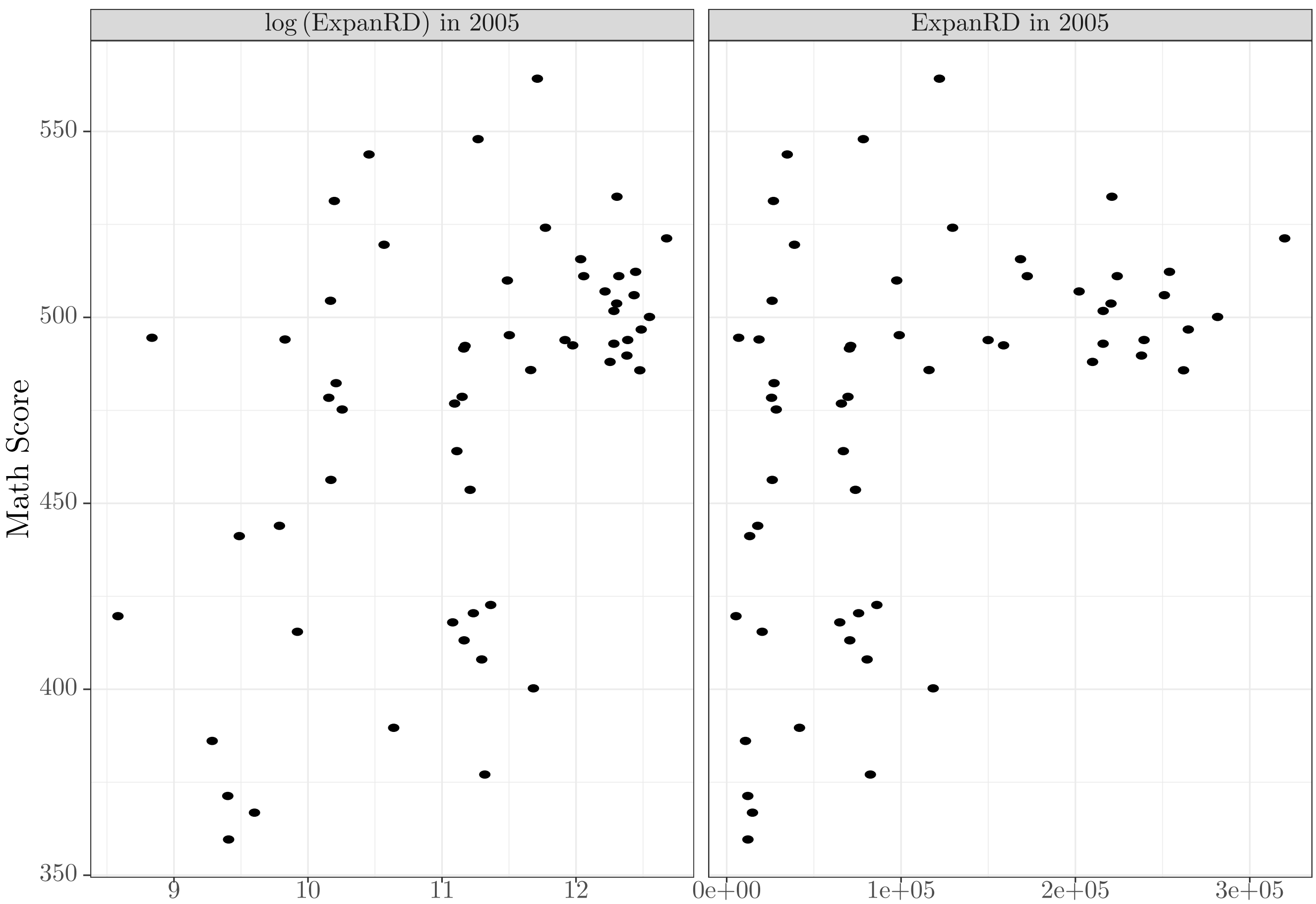

Figure 1: Two scatter plots: (a) log **ExpOneRD** in 2005 × Math Score, (b) **ExpOneRD** in 2005 × Math Score,

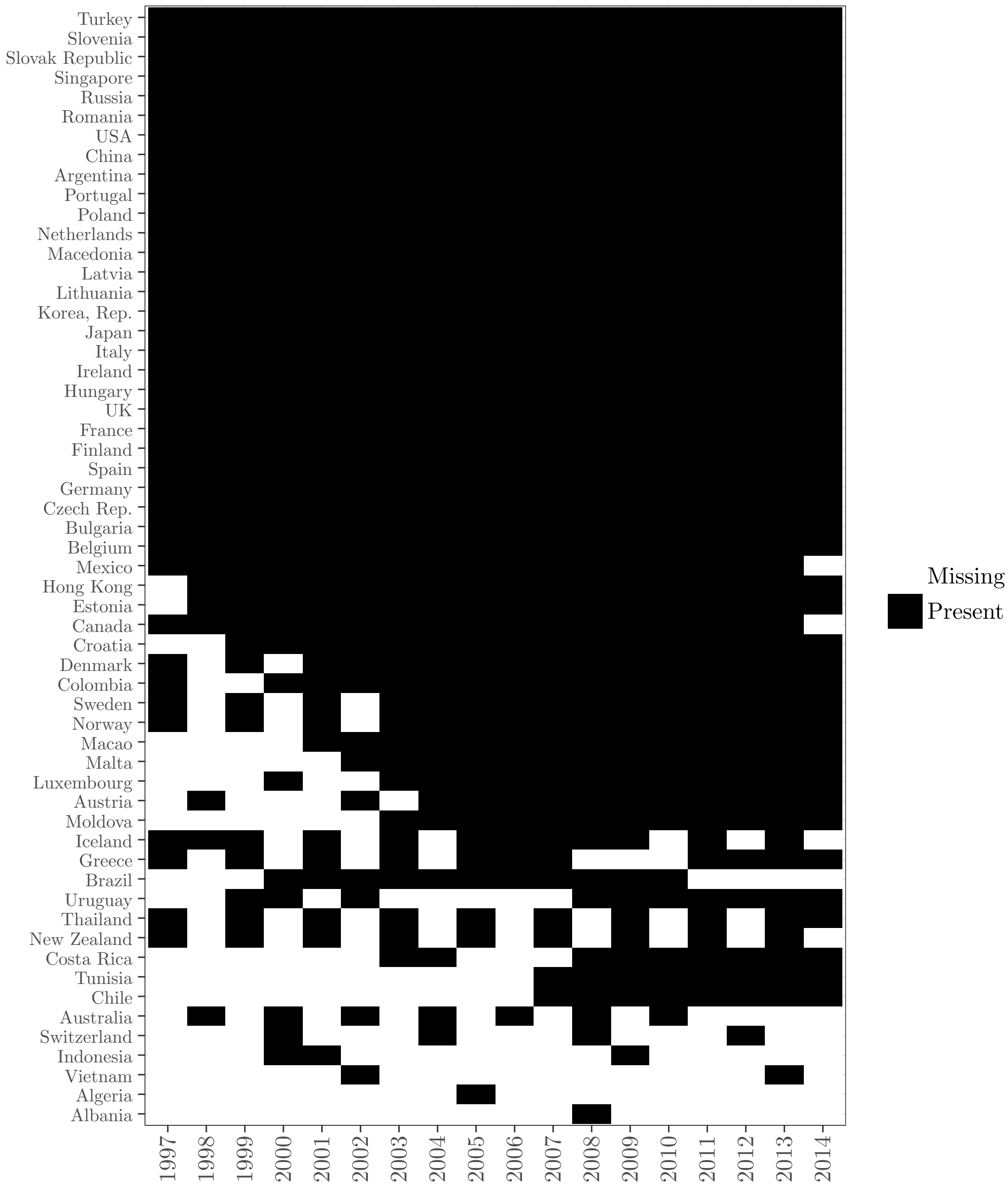

Figure 2: Representation of the missing values in the data $\mathcal{X}$

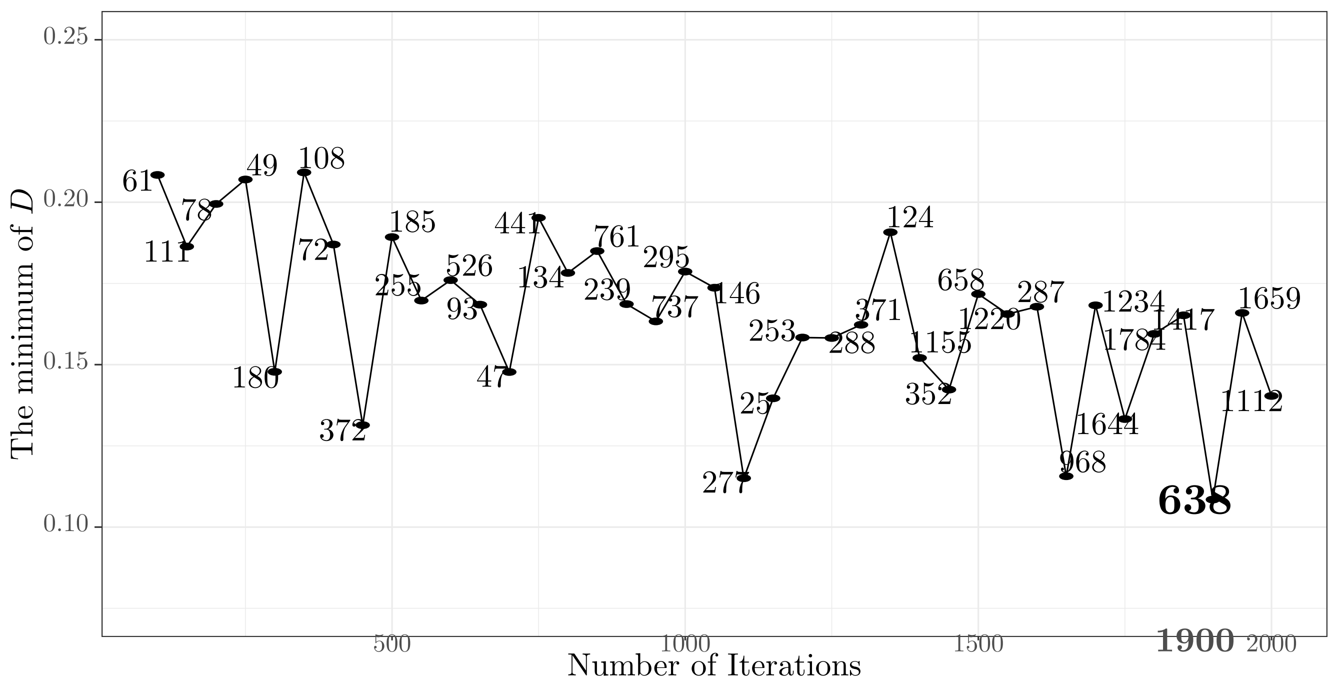

Figure 3: Result of the BNII.

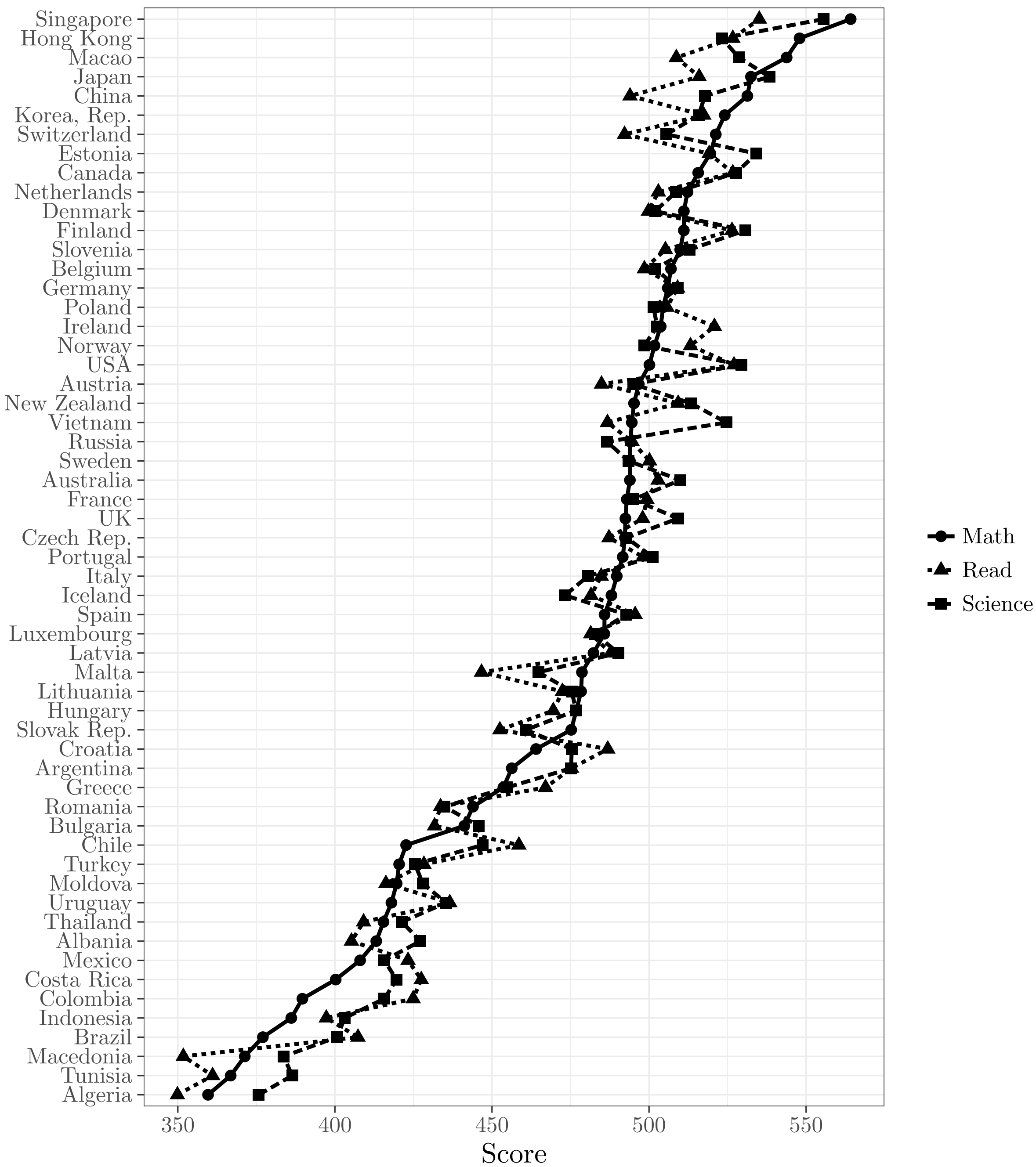

Figure 4: PISA scores

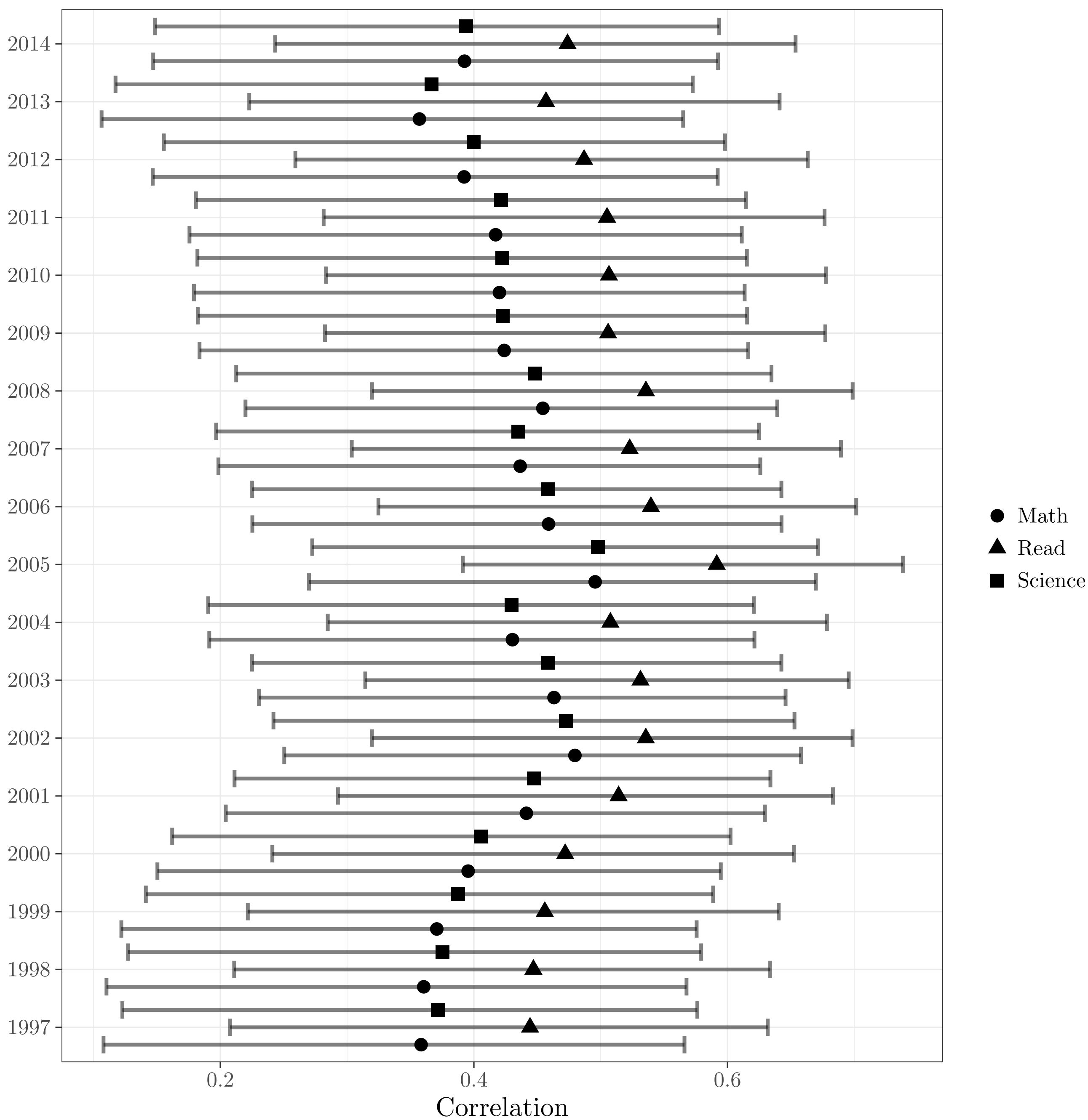

Figure 5: Correlation and their confidence intervals of PISA scores with *Investment* variables from 1997 to 2014

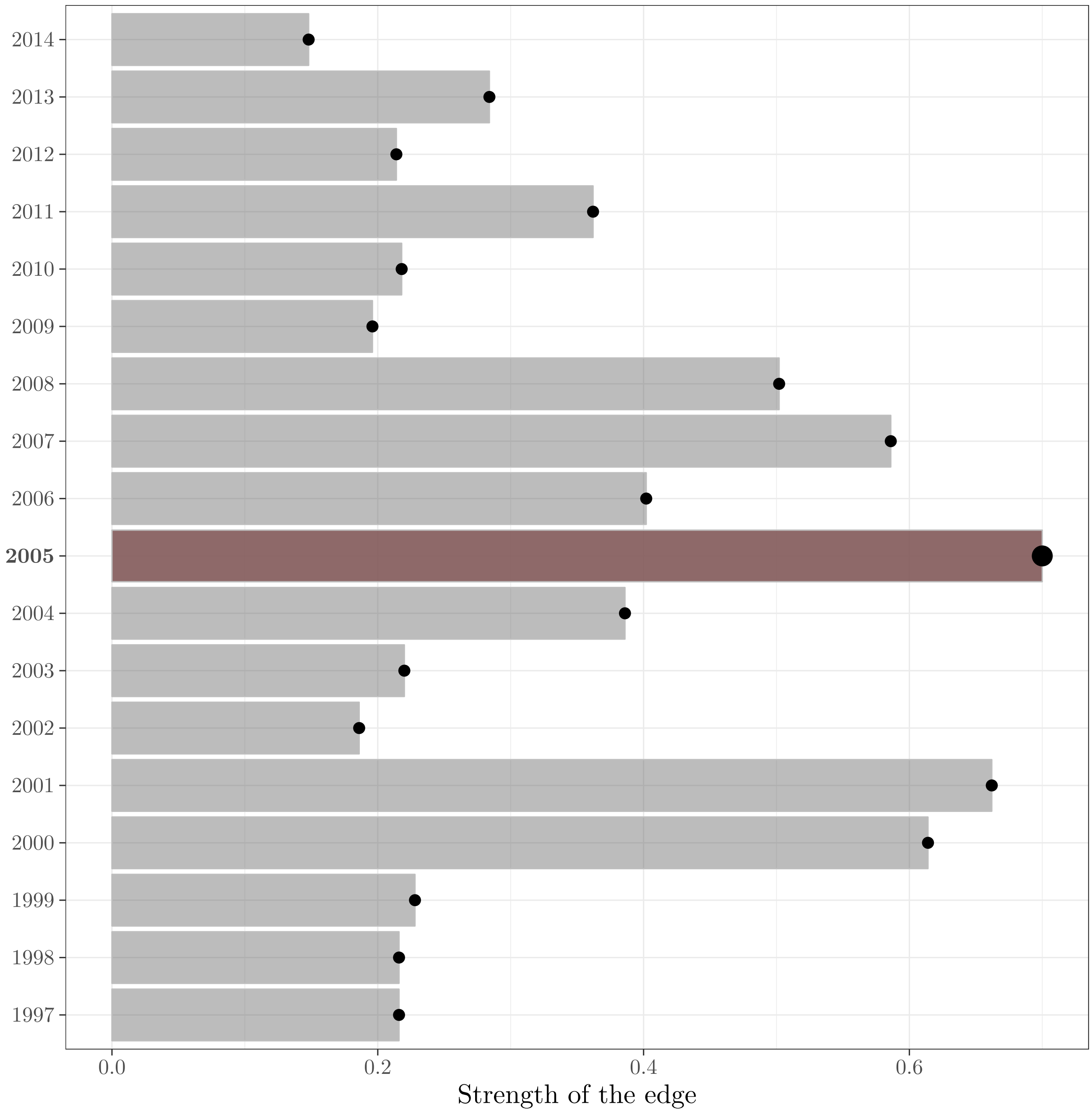

Figure 6: Strength of the link R&D and Reading Score

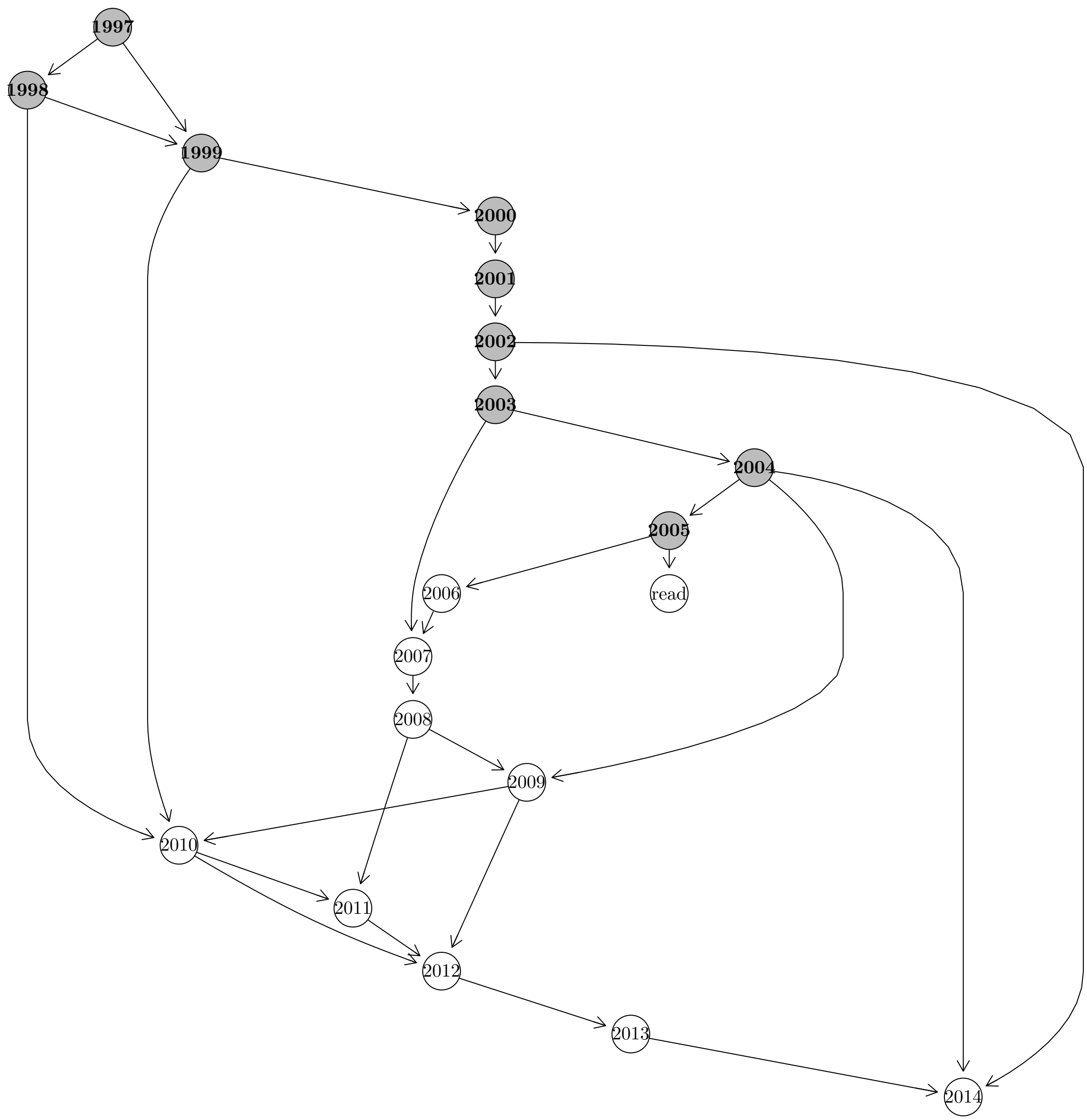

Figure 7: Estimated DBN for Reading Score by highlighting the path to the Math-variable

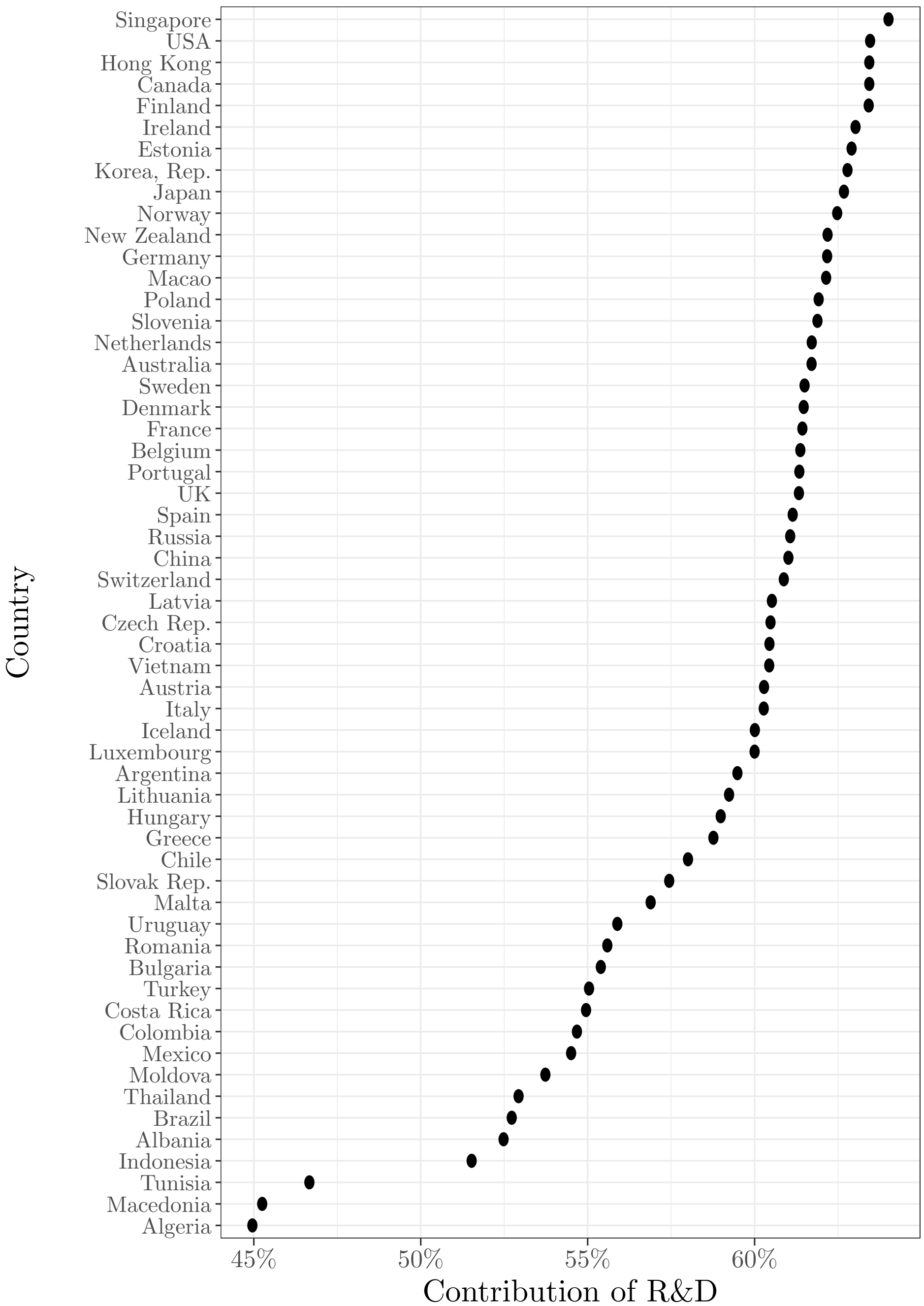

Figure 8: Contribution of R&D in the performance of Education as defined in (10) by country

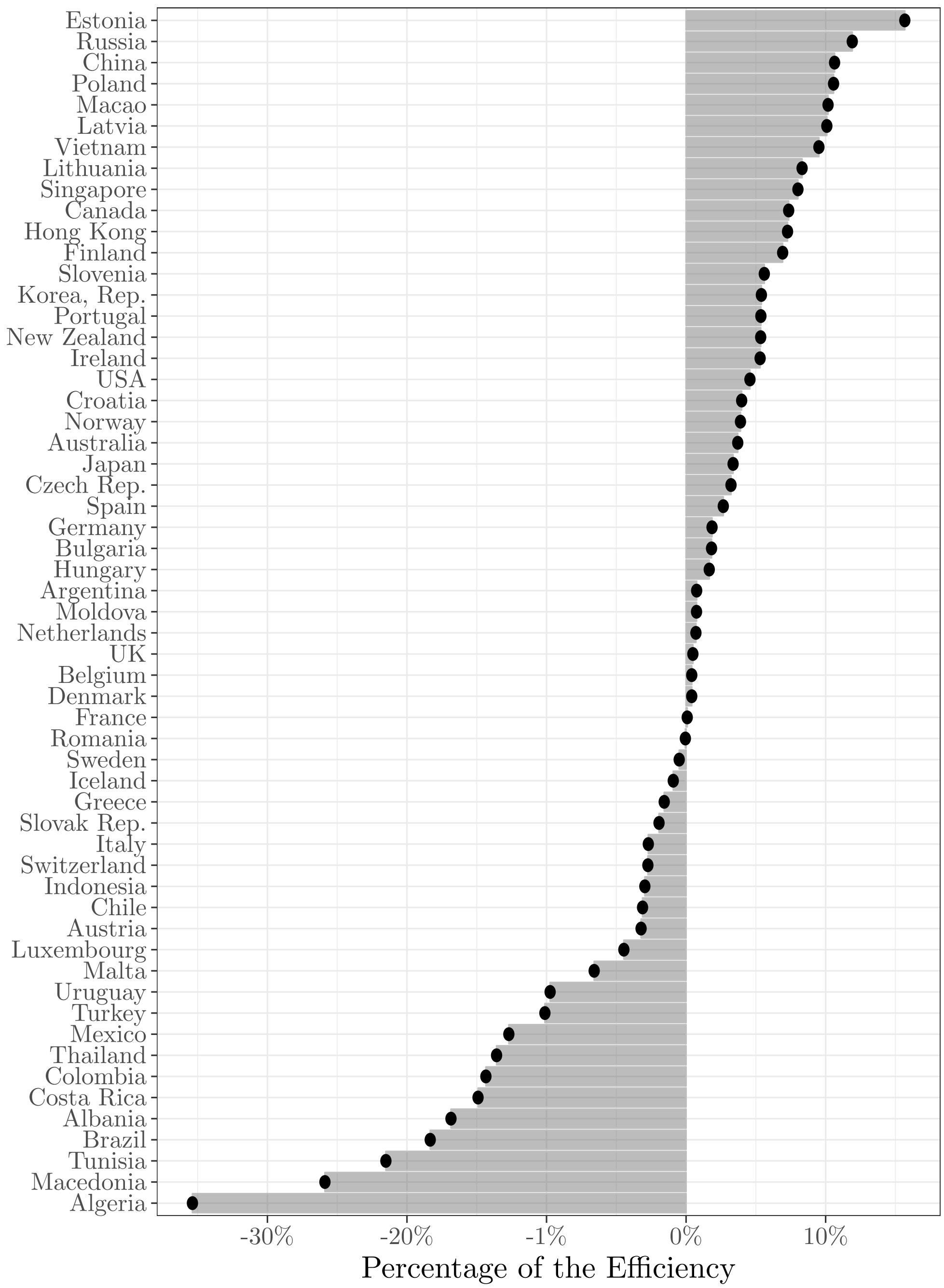

Figure 9: Efficiency Index defined in (11) by country

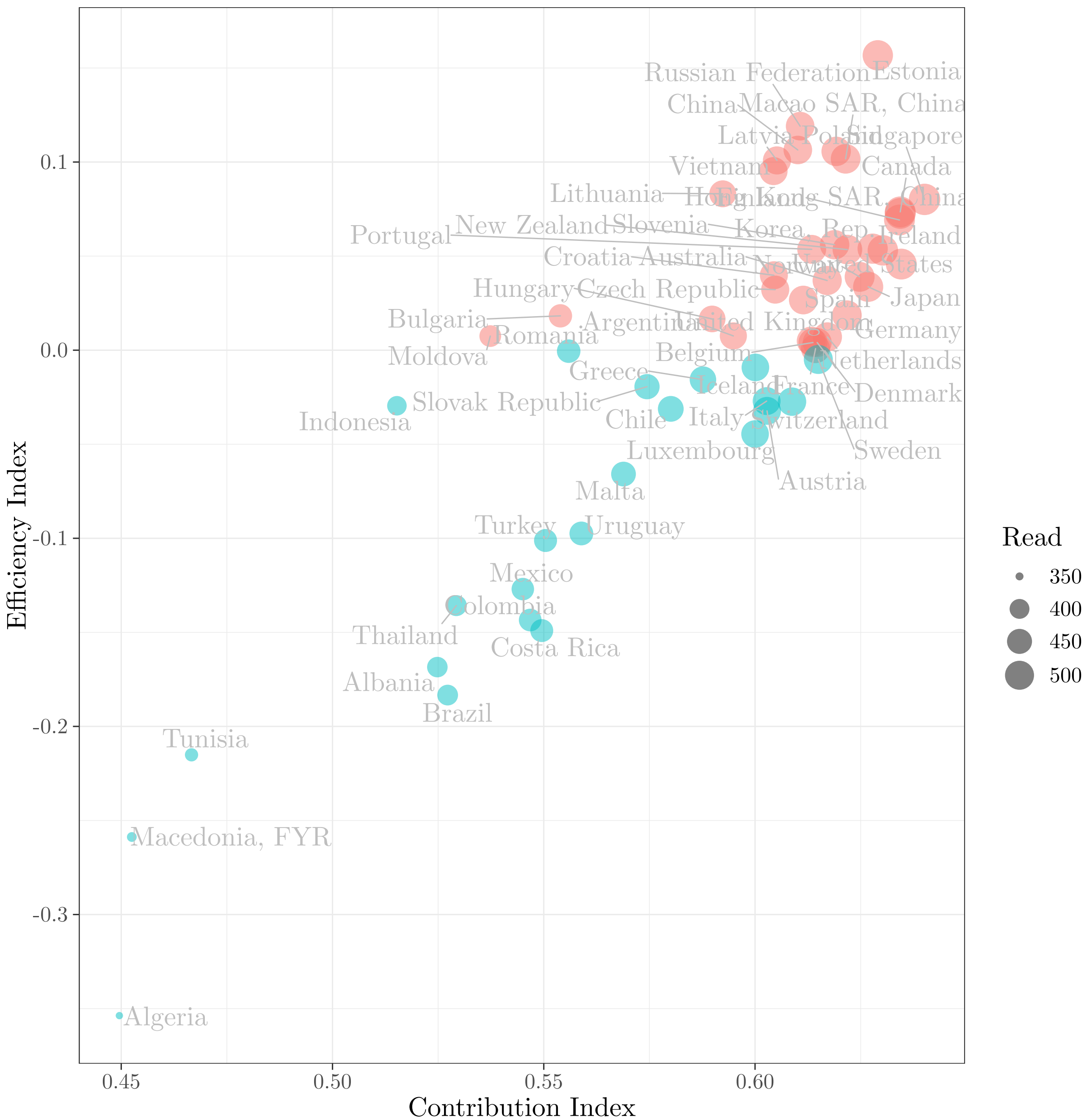

Figure 10: Contribution Index × Efficiency Index × PISA score

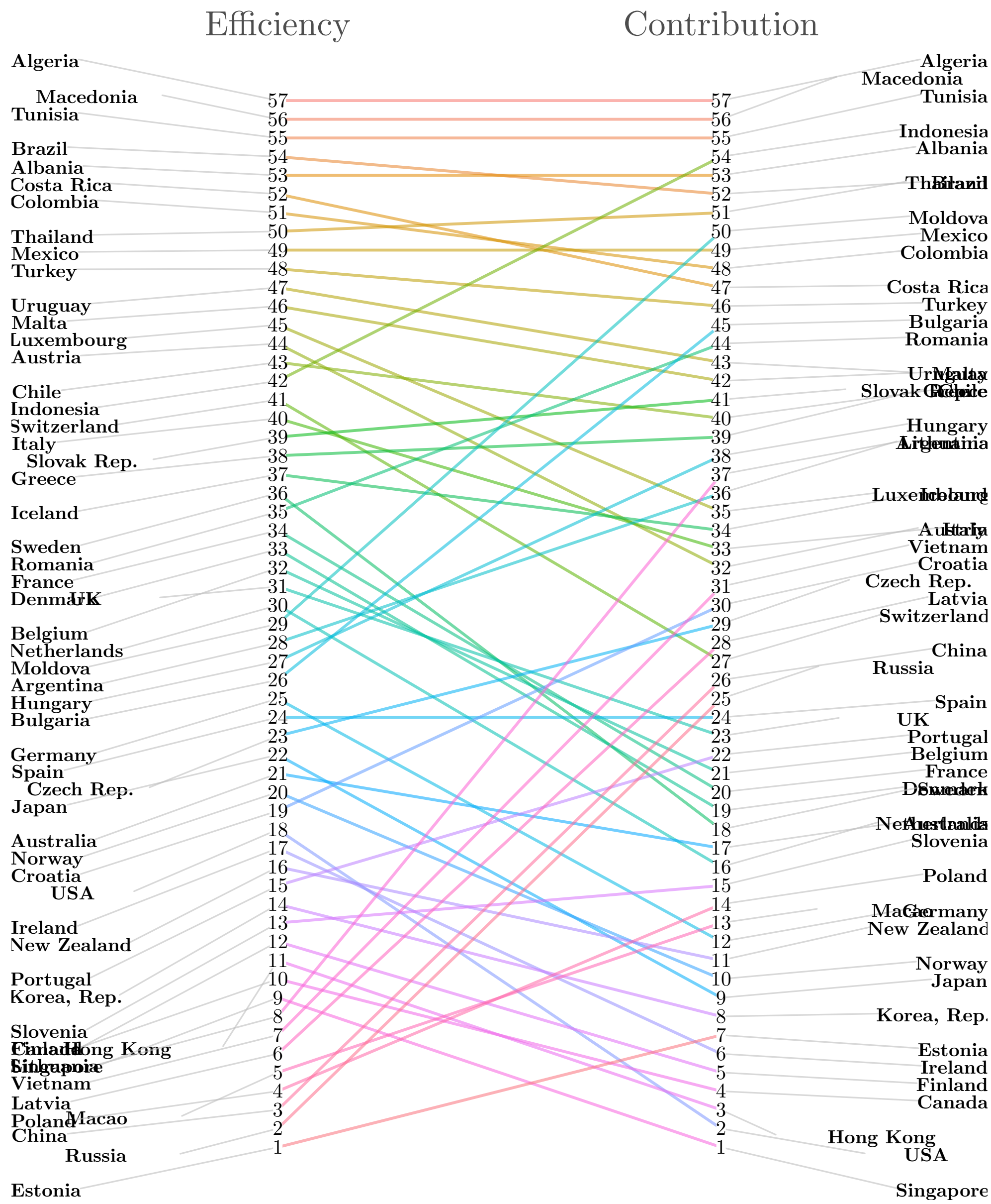

Figure 11: Ranking of countries in terms of Efficiency and Contribution Indices